ISSN 2616-5775 Vol. 9, Issue 9: 15-22, DOI: 10.25236/AJCIS.2026.090903

Driving Risk Identification and Assessment for Freeway Mainline Vehicles Based on Machine Learning

Ruiyang Xue[1,a,*]

[1]*School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, China*
[a]*xry2024200736@my.swjtu.edu.cn*
**Corresponding author*

Abstract: *Freeway traffic combines high speeds with little room for recovery, so a short car-following conflict can escalate into a severe crash within seconds. This paper develops a machine-learning framework that identifies and predicts the driving risk level of individual freeway mainline vehicles from trajectory data. Risk is described by two surrogate safety measures that accumulate over a whole vehicle trajectory instead of being evaluated at a single instant: the time-integrated time-to-collision (TIT), which measures how long a vehicle stays in a hazardous state, and the potential collision severity (PCS), which measures how violent a potential collision would be. Trajectories from the US 101 section of the Next Generation Simulation (NGSIM) dataset are cleaned, restricted to passenger cars in a car-following state, and converted into the two indicators. K-means, hierarchical and spectral clustering are compared for labelling risk levels, and the number of levels is selected with the elbow method, the silhouette coefficient and the Calinski-Harabasz index. K-means with four levels (no, low, medium and high risk) gives the most compact and interpretable partition, and the four levels differ consistently in speed, in spacing and in both risk indicators. The labelled data are then used to train two ensemble classifiers, XGBoost and random forest, on sliding windows extracted from I-80 trajectories. Because the high-risk class is scarce, the training set is re-sampled before learning. Random forest reaches the highest accuracy (98.35%) and F1 score (94.97%), whereas XGBoost reaches the highest overall recall (93.09%) and recovers the largest share of high-risk windows.*

Keywords: *Traffic Engineering, Machine Learning, Driving Risk Assessment, Surrogate Safety Measures, Ensemble Learning*

1. Introduction

Road traffic injury remains one of the most serious public-health problems. In China the number of recorded traffic accidents grew continuously between 2015 and 2020, with about 200,000 people injured and 50,000 killed every year and with direct property losses of roughly one billion yuan per year [1]. Driver-related factors are the dominant cause of freeway crashes and account for about 85.8% of them [2]. Frequent lane changing, speeding and other violations, often combined with an inadequate perception of the surrounding traffic, are among the most common triggers [3,4], and because a driver has only a fraction of a second in which to react, a risk that is not recognised early quickly becomes a crash.

Freeways are designed as fully enclosed, directionally separated and access-controlled roads, so that vehicles can travel at high speed with as little interference as possible. The same design amplifies the consequence of any failure. Relative speeds are high, the energy released in a collision is large and casualties are correspondingly severe. Identifying the driving risk of freeway mainline vehicles before a crash happens is therefore more useful than analysing crashes after they have occurred, both for in-vehicle warning and for traffic management decisions.

Crash records alone cannot support this task well. Crashes are rare and partly random events, so a road segment with no recorded crash may still be dangerous, and models fitted to crash counts are easily biased. Surrogate safety measures (SSMs) avoid this problem by quantifying observable pre-crash behaviour, such as how close two vehicles come in time or in space, and by treating near-miss interactions as a continuous indicator of safety [5]. SSMs have become the main tool for safety evaluation where crash data are scarce [6], and they transfer naturally to the safety modelling of connected and automated vehicles [5].

ISSN 2616-5775 Vol. 9, Issue 9: 15-22, DOI: 10.25236/AJCIS.2026.090903

Machine learning has changed the way in which these measures are used. Instead of relying on fixed thresholds, clustering and classification models learn the boundaries between risk levels directly from data, which makes the assessment more objective and more reproducible [7-10]. Two gaps remain. First, most SSM-based studies describe risk at a single instant, so a vehicle that repeatedly passes through short hazardous states is not distinguished from one that is continuously exposed. Second, most models are trained and tested on the same road section, so it is unclear whether the labelled risk levels transfer to another site.

This paper addresses both gaps. Two surrogate safety measures that accumulate over a whole trajectory, the time-integrated time-to-collision (TIT) and the potential collision severity (PCS), are used together as a two-dimensional description of temporal exposure and of crash severity. A clustering-based labelling procedure is then built and validated, in which K-means, hierarchical and spectral clustering are compared with three internal validity indices, so that the risk levels are derived from the data rather than fixed by assumption. Finally, the resulting labels are transferred to a second freeway site and used to train XGBoost and random forest classifiers on sliding windows of trajectory data, with re-sampling to compensate for the scarcity of high-risk windows [11,12].

The remainder of the paper is organised as follows. Section 2 describes the trajectory data, the surrogate safety measures, the clustering procedure and the ensemble models. Section 3 reports the distributions of the risk indicators, the choice of the number of risk levels, the comparison of the clustering algorithms, the characteristics of the resulting levels and the prediction performance of the two ensemble models. Section 4 summarises the main findings and the limitations of the study.

2. Data and Methods

2.1. Study Data and Preprocessing

The data come from the Next Generation Simulation (NGSIM) programme, which provides vehicle trajectories extracted from high-resolution video at ten frames per second. Two freeway sections are used. The US 101 (Hollywood Freeway) dataset recorded southbound traffic in Los Angeles on 15 June 2005 between 08:05 and 08:20 on a study section about 2100 ft long, with five mainline lanes and an auxiliary lane between the Ventura Boulevard on-ramp and the Cahuenga Boulevard off-ramp; eight cameras covered the section. The I-80 dataset is used as an independent second site, so that the labelled risk levels can be tested on traffic that was not used to build them. NGSIM is known to contain position and speed noise introduced by video processing and has been critically evaluated in the literature [13], so the indicators used here are computed from screened trajectories.

Preprocessing followed five steps. Missing values were filled, outliers were detected and removed, and duplicated records were discarded. Only passenger cars were kept, because motorcycles and trucks differ strongly from cars in size and in acceleration and braking performance and form only a small part of the data. Speeds, accelerations, spacings and positions were converted from imperial to metric units. Vehicles in car-following or lane-changing states were retained by keeping only the records that have a valid preceding vehicle, because longitudinal conflicts arise in these states. Finally, only the fields needed by the risk indicators were kept and the indicator values were computed for every frame.

For the prediction task the aggregated per-vehicle description used for labelling is replaced by a time-resolved description. Each trajectory is divided into sliding windows of 200 frames that advance by 20 frames, so that a window covers 20 s of driving and consecutive windows overlap by 90%. The risk indicators are computed inside every window and the windows of a vehicle are ordered in time. The first 70% of the windows of each vehicle form the training set and the remaining 30% form the test set, which keeps the evaluation chronological and prevents future information from leaking into training.

2.2. Surrogate Safety Measures

Surrogate safety measures can be grouped by the quantity they describe: temporal proximity, spatial proximity and the deceleration required to avoid a crash [5,7]. The most widely used measure of temporal proximity is the time-to-collision (TTC). For a vehicle i that follows a vehicle i-1, the TTC is the time that remains before a collision if both vehicles keep their current speeds [14]:

$$TTC_i(t) = \frac{d_i(t)}{\Delta v_i(t)} \tag{1}$$

where $d_i(t)$ is the gap between the front of the follower and the rear of the leader and $\Delta v_i(t)$ =

ISSN 2616-5775 Vol. 9, Issue 9: 15-22, DOI: 10.25236/AJCIS.2026.090903

$v_i(t) - v_{i-1}(t)$ is the speed difference. A collision can occur only while the follower is faster than the leader, so the TTC is treated as infinite otherwise. Because a single TTC value describes one instant, a threshold based on TTC alone cannot distinguish a vehicle that is briefly close from one that stays close for many seconds. The time-integrated time-to-collision (TIT) removes this weakness by accumulating only the time that a vehicle spends below a critical threshold [15]:

$$TIT_i = \sum_{t=0}^{N} (TTC^* - TTC_i(t))\tau_{sc} \quad (2)$$

where TTC^* is the threshold and τ_{sc} = 0.1 s is the length of one time interval, so that the index sums the excess risk over the observation period. The threshold was set to TTC^* = 4 s, the value usually recommended for this purpose, because it keeps the indicator sensitive to emerging risk while still ignoring ordinary car-following interactions [15]. Larger TIT values therefore mean longer exposure.

Severity is described by the potential collision severity (PCS), which integrates the speed difference, the distance between the vehicles and the angle of the interaction over the trajectory [16]:

$$PCS = \int \frac{\Delta v^2}{d} f(\theta) dt \quad (3)$$

where Δv is the speed difference between the two vehicles, d is the distance between them, $f(\theta)$ is a function that weights the influence of the interaction angle θ, and the integral is taken over the duration of the trajectory. Unlike an instantaneous measure, PCS describes the overall severity of the risk that a vehicle carries throughout the observation period.

TIT and PCS were selected because they describe the two dimensions that matter for risk, exposure and severity, and because both of them accumulate over a trajectory. Measures of spatial proximity such as the proportion of stopping sight distance, and deceleration-based measures such as the deceleration rate to avoid a crash, were not used here. Deceleration-based measures depend on an assumed braking capability and on the reaction time of the driver, both of which are difficult to observe, whereas the time-based and severity-based measures used here need only the gap and the speeds of the two vehicles [17-21].

2.3. Risk-Level Labelling by Clustering

No crash label is available for an individual trajectory, so the risk levels are defined without supervision. Every trajectory is represented by its pair of standardised indicator values, (TIT, PCS), and three clustering algorithms are compared.

K-means is used as the reference method. It partitions the n trajectories into k clusters by minimising the sum of squared distances between every point and the centre of its own cluster:

$$SSE = \sum_{i=1}^{k} \sum_{p \in C_i} (p - m_i)^2 \quad (4)$$

The algorithm alternates between assigning each point to the nearest centre and recomputing the centres, and it was restarted from several random initialisations to reduce the effect of local optima. Hierarchical clustering merges the two closest clusters repeatedly and therefore does not need k in advance, but its cost grows quickly with the number of trajectories. Spectral clustering embeds the data in the eigenspace of an affinity matrix before K-means is applied, which allows clusters that are not convex to be separated.

The number of levels was chosen with three internal indices. The elbow method plots SSE against k and looks for the value after which an additional cluster gives little further improvement. The silhouette coefficient compares the mean distance from a point to the other points of its own cluster with the mean distance to the points of the nearest other cluster; it lies between -1 and 1 and values close to 1 indicate compact and well-separated clusters:

$$s(i) = \frac{b(i) - a(i)}{max(a(i), b(i))} \quad (5)$$

where a(i) is the mean distance between point i and the other points of its own cluster and b(i) is the smallest mean distance between point i and the points of any other cluster. The Calinski-Harabasz index is the ratio of between-cluster to within-cluster dispersion, and larger values are better:

$$CH(k) = \frac{\frac{tr(B_k)}{k-1}}{\frac{tr(W_k)}{m-k}} \quad (6)$$

ISSN 2616-5775 Vol. 9, Issue 9: 15-22, DOI: 10.25236/AJCIS.2026.090903

where B_k and W_k are the between-cluster and within-cluster covariance matrices of the partition into k clusters, m is the number of trajectories and tr denotes the trace of a matrix. The three indices were computed for k from 1 to 10.

2.4. Ensemble Models and Evaluation

Two ensemble classifiers were used. XGBoost grows a sequence of regression trees in which every new tree corrects the residual error of the trees that have already been grown, and it minimises a regularised objective [22]:

$$Obj = \sum_{i=1}^{n} L(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{7}$$

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{8}$$

where L is the loss function, y_i and $\hat{y}_i$ are the observed and the predicted label, K is the number of trees, T is the number of leaves of a tree, w_j is the weight of leaf j, and γ and λ are regularisation parameters that penalise the number of leaves and the leaf weights. Random forest instead grows many independent decision trees on bootstrap samples of the training set, and at each split only a random subset of the features is considered [23]. Voting over the trees reduces the variance of the prediction and makes the classifier insensitive to outliers.

Four metrics were used to compare the predicted labels with those obtained by clustering. Accuracy measures the overall proportion of correct predictions:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

$$F1 = \frac{2PR}{P + R} \tag{12}$$

where TP, TN, FP and FN are the numbers of true positives, true negatives, false positives and false negatives, and P and R denote precision and recall. Precision describes how many of the windows predicted as high-risk really are high-risk, recall describes how many of the truly high-risk windows are found, and the F1 score is the harmonic mean of the two.

The four risk levels are strongly imbalanced; in the US 101 data the high-risk level contains only 200 of the 5810 trajectories. A classifier trained on such data tends to predict the majority level and easily overfits the few high-risk windows. The training set was therefore re-sampled before learning. Synthetic minority over-sampling was used to generate new high-risk samples by interpolating between a high-risk window and one of its nearest high-risk neighbours [24,25], and the majority levels were randomly under-sampled. Re-sampling was applied to the training set only, so that the reported metrics describe the original class distribution.

3. Results and Discussion

3.1. Distribution of the Risk Indicators

Figure 1 shows the distributions of TIT and PCS over the US 101 trajectories. Both indicators accumulate over time, so larger values mean longer exposure or greater potential severity. Most vehicles have a TIT below 8 and a PCS below 20, and only a small group exceeds these ranges. The strongly right-skewed shape is consistent with the traffic conditions on the section, where most vehicles travel in a stable car-following state and only a minority spend a sustained period in a hazardous state or approach another vehicle with a large relative speed.

ISSN 2616-5775 Vol. 9, Issue 9: 15-22, DOI: 10.25236/AJCIS.2026.090903

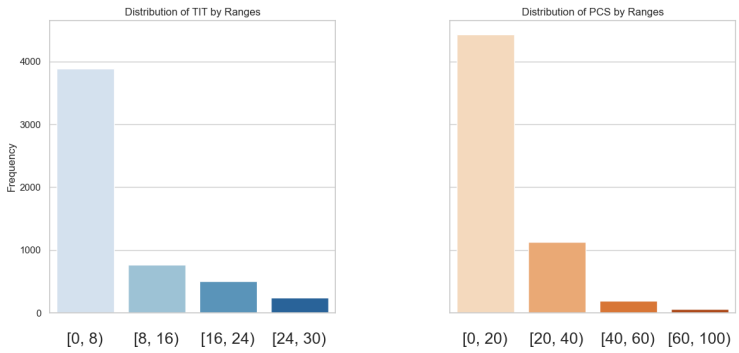


Figure 1: Distributions of the accumulated time-to-collision (TIT) and of the potential collision severity (PCS).

3.2. Clustering-Based Risk Labelling

The elbow curve in Figure 2 falls steeply up to three clusters and flattens afterwards, and the largest reduction in curvature occurs at k = 4. The silhouette coefficient and the Calinski-Harabasz index lead to the same choice: four levels separate the trajectories better than three or five, so the trajectories were labelled as no risk, low risk, medium risk and high risk.

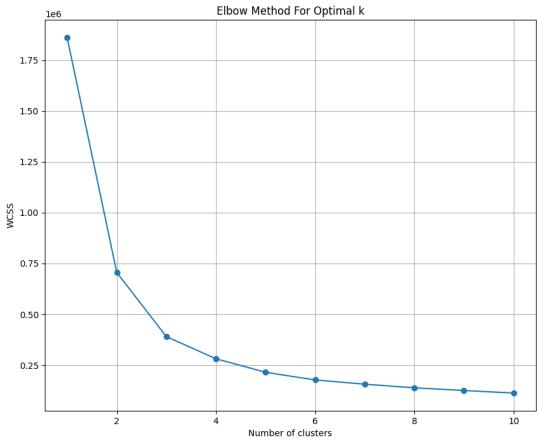


Figure 2: Sum of squared errors as a function of the number of clusters (elbow curve).

Table 1 compares the three clustering algorithms. K-means assigns 3444 trajectories to the no-risk level, 1439 to the low-risk level, 727 to the medium-risk level and 200 to the high-risk level, and it achieves both the highest Calinski-Harabasz score (10573) and the highest silhouette coefficient (0.54). Hierarchical clustering produces a similar shape but a lower silhouette coefficient (0.50), and its cost is much higher because clusters must be merged repeatedly. Spectral clustering performs worst on this data set (silhouette 0.28): it places 2244 trajectories in the medium-risk level and 1512 in the high-risk level, and leaves only 728 trajectories in the no-risk level. The weak result is consistent with the strongly skewed indicator values, because spectral clustering relies on an affinity graph in which the dense no-risk region dominates.

Table 1: Comparison of the three clustering algorithms on the US 101 data.

Algorithm	No risk	Low risk	Medium risk	High risk	CH index	Silhouette
K-means	3444	1439	727	200	10573	0.54
Hierarchical	3773	1273	509	255	9212	0.50
Spectral	728	1326	2244	1512	3636	0.28

Figure 3 shows the same comparison as proportions. K-means gives the most even progression from the no-risk level (59.3%) to the high-risk level (3.4%), which is the behaviour expected of a scale that is to be used for warning. Hierarchical clustering is even more concentrated in the no-risk level (64.9%) and leaves only 4.4% of the trajectories in the high-risk level, so the few high-risk cases are harder to learn. Spectral clustering inverts the ordering and places 26.0% of the trajectories in the high-risk level, which is not consistent with the indicator distributions shown in Figure 1.

ISSN 2616-5775 Vol. 9, Issue 9: 15-22, DOI: 10.25236/AJCIS.2026.090903

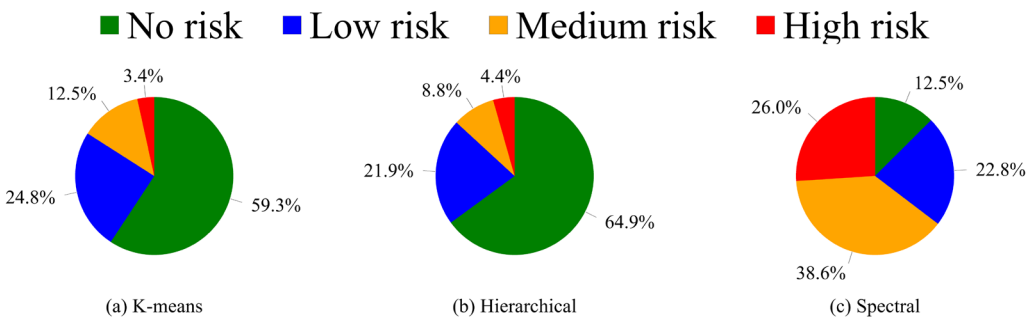


Figure 3: Risk counts produced by (a) K-means, (b) hierarchical and (c) spectral clustering.

3.3. Characteristics of the Four Risk Levels

Table 2 gives the mean speed, acceleration, spacing and indicator values of the four levels obtained by K-means. The levels are ordered and internally consistent. Vehicles in the no-risk level travel fastest (mean speed 4.43), keep the largest spacing (9.47) and have very small TIT (1.13) and PCS (7.54) values. As the risk level rises, the indicator values increase while speed and spacing fall: the high-risk level has the lowest mean speed (2.65) and the smallest mean spacing (7.65), together with by far the largest TIT (56.95) and PCS (52.26).

Table 2: Mean feature values of the four risk levels obtained by K-means clustering.

Risk level	Speed	Acceleration	Spacing	TIT	PCS
No risk	4.43	0.031	9.47	1.13	7.54
Low risk	3.26	0.014	8.13	10.95	18.40
Medium risk	2.91	0.013	7.85	26.96	30.82
High risk	2.65	0.014	7.65	56.95	52.26

The ordering also has a behavioural interpretation. Vehicles in the no-risk level keep a comfortable distance and hardly change speed, so neither indicator accumulates. Low-risk vehicles show a moderate exposure but a small severity, which corresponds to a short interaction that is quickly resolved. Medium-risk vehicles combine a rising exposure with a rising severity, and high-risk vehicles reach the largest values of both indicators with the largest spread, which means that they are not a homogeneous group but contain a small number of strongly hazardous trajectories. Because both indicators accumulate over the trajectory, a vehicle cannot reach the high-risk level through a single extreme frame; it has to be exposed repeatedly.

3.4. Prediction Performance of the Ensemble Models

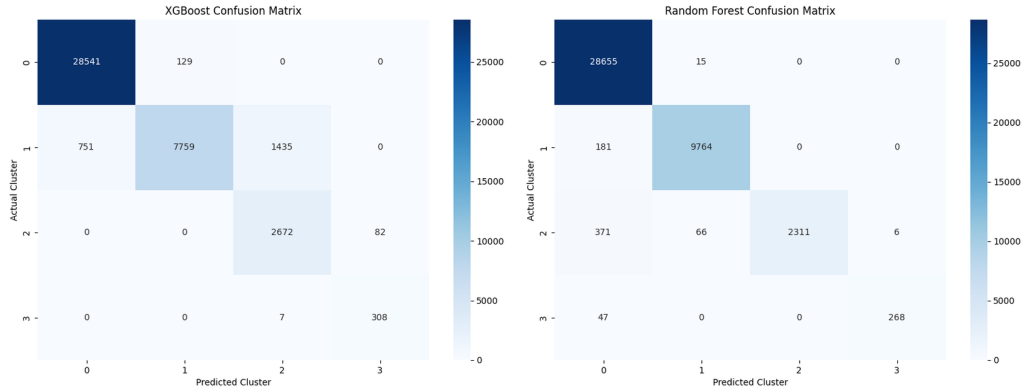
The labels obtained on US 101 were transferred to the I-80 sliding windows. Within every window the same two indicators were computed, K-means was applied to the training windows in order to assign labels, and the test windows were then classified by XGBoost and by random forest. Table 3 and Figure 4 summarise the result.

Table 3: Prediction performance of the two ensemble models on the test set.

Model	Accuracy	Precision	Recall	F1 score
XGBoost	94.23%	84.93%	93.09%	87.67%
Random forest	98.35%	98.73%	91.78%	94.97%

Random forest is the stronger model overall. It reaches an accuracy of 98.35%, a precision of 98.73% and an F1 score of 94.97%, and its confusion matrix in Figure 4(b) is very nearly diagonal, with only a few windows misclassified from each actual level. XGBoost reaches a lower accuracy (94.23%) and a lower precision (84.93%), because many low-risk windows are classified as medium risk, which appears as the large off-diagonal cell in Figure 4(a).

ISSN 2616-5775 Vol. 9, Issue 9: 15-22, DOI: 10.25236/AJCIS.2026.090903



(a) XGBoost (b) Random forest

Figure 4: Confusion matrices of (a) XGBoost and (b) random forest on the test set.

The picture changes when recall is considered. XGBoost obtains the higher overall recall (93.09% against 91.78%) and recovers 308 of the 315 high-risk windows of the test set, whereas random forest recovers 268 of them. Random forest is therefore the more conservative model: it rarely raises a false alarm, but it also misses more genuinely hazardous cases. XGBoost is more willing to assign a window to a higher risk level and consequently misses fewer of them. For a warning application the second behaviour is preferable, because a missed high-risk window is more costly than a false alarm, and the re-sampling of the high-risk class is what makes that behaviour possible.

Two further observations follow. The first is that re-sampling mattered more than the choice of ensemble algorithm: before re-sampling, both models concentrated their errors on the small high-risk class, and the improvement after re-sampling was concentrated in the same class. The second is that the risk scale is transferable. K-means was applied to the I-80 windows without any change of parameters, and the resulting labels were then predicted with an accuracy above 94%, which suggests that a scale built on accumulated TIT and PCS values describes freeway car-following risk beyond the section on which it was defined.

4. Conclusions

This paper developed and tested a machine-learning framework for identifying and predicting the driving risk level of freeway mainline vehicles. Two surrogate safety measures that accumulate over a trajectory, the time-integrated time-to-collision (TIT) and the potential collision severity (PCS), were used to describe the temporal exposure and the severity of each vehicle. On the US 101 data the elbow method, the silhouette coefficient and the Calinski-Harabasz index agreed that four levels give the best separation, and K-means produced the most compact and even partition, with a silhouette coefficient of 0.54 against 0.50 for hierarchical clustering and 0.28 for spectral clustering. The four levels are ordered in the expected way: mean speed and mean spacing decrease from the no-risk to the high-risk level while both indicators increase, and the high-risk level has both the largest indicator values and the largest spread.

The labels were then transferred to sliding windows of I-80 trajectories and predicted by two ensemble classifiers after the training set had been re-sampled. Random forest achieved the best overall performance (98.35% accuracy and an F1 score of 94.97%), while XGBoost achieved the highest overall recall (93.09%) and recovered the largest number of high-risk windows. This division of labour is useful in practice: if the model is used to raise warnings, XGBoost is the safer choice, whereas if it is used to filter candidates for further inspection, random forest is more accurate.

Three limitations remain. Surrogate measures describe pre-crash behaviour but cannot reproduce the relation between that behaviour and crashes that actually occur, so the labels should be validated once crash data become available. Only passenger cars were modelled, whereas conflicts between vehicles of different classes are an important part of real freeway risk. Finally, the trajectories come from two US freeways whose vehicle dynamics and road design differ from those of Chinese freeways, so the transferability of the scale to domestic data still has to be tested.

Acknowledgements

The authors thank the Federal Highway Administration for making the NGSIM trajectory data

ISSN 2616-5775 Vol. 9, Issue 9: 15-22, DOI: 10.25236/AJCIS.2026.090903

publicly available, and Southwest Jiaotong University for supporting this undergraduate research project.

References

[1] China Industry Information Network (2020) Analysis of Car Ownership and the Number of Drivers in China in 2020. (In Chinese) http://www.chyxx.com
[2] Zhang, B. (2019) Analysis of Freeway Traffic Accident Patterns and Influencing Factors. Master's Thesis, People's Public Security University of China, Beijing. (In Chinese)
[3] Reason, J., Manstead, A., Stradling, S., Baxter, J. and Campbell, K. (1990) Errors and Violations on the Roads: A Real Distinction? Ergonomics, 33, 1315-1332.
[4] Wang, X. and Xu, X. (2019) Assessing the Relationship between Self-Reported Driving Behaviors and Driver Risk Using a Naturalistic Driving Study. Accident Analysis & Prevention, 128, 8-16.
[5] Wang, C., Xie, Y., Huang, H. and Liu, P. (2021) A Review of Surrogate Safety Measures and Their Applications in Connected and Automated Vehicles Safety Modeling. Accident Analysis & Prevention, 157, 106157.
[6] Wang, Y., Li, Z., Liu, P., et al. (2024) Surrogate Safety Measures for Traffic Oscillations Based on Empirical Vehicle Trajectories Prior to Crashes. Transportation Research Part C: Emerging Technologies, 161, 104543.
[7] Shi, X., Wong, Y.D., Li, M.Z.F., et al. (2018) Key Risk Indicators for Accident Assessment Conditioned on Pre-Crash Vehicle Trajectory. Accident Analysis & Prevention, 117, 346-356.
[8] Chen, T., Shi, X., Wong, Y.D., et al. (2020) Predicting Lane-Changing Risk Level Based on Vehicles' Space-Series Features: A Pre-Emptive Learning Approach. Transportation Research Part C: Emerging Technologies, 116, 102646.
[9] Xue, Q., Jiang, Y. and Lu, J. (2020) Identification Method of Dangerous Driving Behaviour Based on Trajectory Data. China Journal of Highway and Transport, 33, 11. (In Chinese)
[10] Zheng, K. (2019) Research on Traffic Operation Risk Prediction Method of Urban Bridge-Tunnel Connection Section Based on Multi-Source Data. Master's Thesis, Chongqing Jiaotong University, Chongqing. (In Chinese)
[11] Shi, X., Wong, Y.D., Li, M.Z.F., et al. (2019) A Feature Learning Approach Based on XGBoost for Driving Assessment and Risk Prediction. Accident Analysis & Prevention, 129, 170-179.
[12] Xing, L., He, J., Li, Y., et al. (2020) Comparison of Different Models for Evaluating Vehicle Collision Risks at Upstream Diverging Area of Toll Plaza. Accident Analysis & Prevention, 135, 105343.
[13] Coifman, B. and Li, L. (2017) A Critical Evaluation of the Next Generation Simulation (NGSIM) Vehicle Trajectory Dataset. Transportation Research Part B: Methodological, 105, 362-377.
[14] Hayward, J. (1971) Near Misses as a Measure of Safety at Urban Intersections. Pennsylvania Transportation and Traffic Safety Center, University Park.
[15] Minderhoud, M.M. and Bovy, P.H.L. (2001) Extended Time-to-Collision Measures for Road Traffic Safety Assessment. Accident Analysis & Prevention, 33, 89-97.
[16] Tang, S., Lu, Y., Liao, Y., Cheng, K. and Zou, Y. (2024) A New Surrogate Safety Measure Considering Temporal-Spatial Proximity and Severity of Potential Collisions. Applied Sciences, 14, 2711.
[17] Cunto, F. and Saccomanno, F.F. (2008) Calibration and Validation of Simulated Vehicle Safety Performance at Signalized Intersections. Accident Analysis & Prevention, 40, 1171-1179.
[18] Meng, X., Xu, H., Wang, H., et al. (2012) Rear-End Conflict in Freeway Work Zones Based on TTC and DRAC. Traffic Information and Safety, 30, 5. (In Chinese)
[19] Oh, C. and Kim, T. (2010) Estimation of Rear-End Crash Potential Using Vehicle Trajectory Data. Accident Analysis & Prevention, 42, 1888-1893.
[20] Xie, J., Qin, Y., Peng, B., et al. (2021) Risk Discrimination and Conflict Prediction of Vehicle Car-Following Behaviour in Multi-Lane Weaving Sections. Journal of Transportation Systems Engineering and Information Technology, 21, 131. (In Chinese)
[21] Wang, B., Gao, L. and Juan, Z. (2018) Conflict Analysis of Lane Changing in Weaving Areas Based on Time-to-Avoidance. Journal of System Simulation, 30, 9. (In Chinese)
[22] Chen, T., He, T., Benesty, M., et al. (2015) XGBoost: Extreme Gradient Boosting. R Package Version 0.4-2, 1-4.
[23] Speiser, J.L., Miller, M.E., Tooze, J. and Ip, E. (2019) A Comparison of Random Forest Variable Selection Methods for Classification Prediction Modeling. Expert Systems with Applications, 134, 93-101.
[24] Chawla, N.V., Bowyer, K.W., Hall, L.O. and Kegelmeyer, W.P. (2002) SMOTE: Synthetic Minority Over-Sampling Technique. Journal of Artificial Intelligence Research, 16, 321-357.
[25] Han, H., Wang, W.Y. and Mao, B.H. (2005) Borderline-SMOTE: A New Over-Sampling Method in Imbalanced Data Sets Learning. Advances in Intelligent Computing, ICIC 2005, Hefei, 878-887.